Ling 315 Lecture notes. 15 January 2007.

1. The big picture. Our aim in this class is to build a system of rules and principles that generates *all* and *only* the sentences of English that are judged grammatical by native speakers. We will call this system a *grammar*.

Our current grammar, given below, consists of a lexicon, which specifies the lexical category of each word, and a set of phrase structure rules, which combine words into phrases, and phrases into larger phrases, the largest phrase being the sentence.

Lexicon

N → *kittens*, *squirrels*, *bed*, …
D → *the*, *those*, …
V → *love*, *hugged*, *slept*, *yawned*, …
P → *on*, *near*, …
A → *happy*, *nice,* …

Phrase Structure Rules

S → NP VP
NP → (D) (A) N (PP)
VP → V (NP) (PP)
PP → P (NP)

Does our current grammar generate *all* sentences of English?
Does our current grammar generate *only* sentences of English?

2. Lexical categories. Certain classes of words behave similarly syntactically (that is, they have the same distribution within a sentence) or morphologically (that is, they combine with the same set of affixes), and this motivates the classification of words in the lexicon into *lexical categories*. For example, the category *noun* (N) represents a class of words that (i) may follow a determiner (D) and an adjective (A) and (ii) may be pluralized.

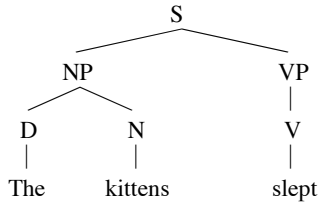
Exercise 1. What are the syntactic categories of the following words? Use the tests for identifying a word's category to support your answer.

real, *hand*, *your, some*, *hurt*, *on, can*, *fool*, *foolish*, *really*, *dream*, *over*

2. Constituent structure. According to our grammar, the individual words in a sentence are organized into groupings, which are themselves organized into larger groupings, the largest grouping being the sentence. These groupings within a sentence are called *constituents*, or *phrases*.

In grouping words into constituents, we are assigning internal structure to sentences. This structure can be represented in a tree diagram. For example, our current grammar generates the following tree:

(1)



Tree terminology. A tree is said to consist of *nodes*, connected by *branches*. The nodes at the bottom of the tree are referred to as *terminal* nodes. A node right above another node is said to be its *mother* and *immediately dominate* it. A node right below another on a branch is said to be its *daughter*. Two daughters of the same mother node are referred to as *sisters*.

Exercise 2: Draw the phrase structure trees for the following sentences.

(2) The kittens slept at midnight.
(3) The kittens from next-door slept at midnight.
(4) Some people put those things on your bed.
(5) Some people put those things on your bed on the floor.
(6) That guy from next-door baked a cake.

3. Motivating constituent structure. While it is natural to rely on intuition in analyzing the structure of sentences, we want to make sure we have good *justification* for positing the structures that we posit. So far, we have justified saying that a given string is a constituent using three types of arguments:

Argument 1: Substitution classes. By positing phrases, we account for the fact that certain sequences of words behave similarly syntactically, that is, form substitution classes. For example, *squirrels* (N), *nice squirrels* (A N), and *the nice squirrels* (D A N) may all occur in the same position within a sentence, e.g., before or after a verb. Thus, we can generalize and say that these form a class, which we refer to as NP.

Argument 2: Constituency tests. Certain rules of English need to make reference to constituents. For example, (for the most part), only constituents may be coordinated, and only constituents may be replaced by pro-forms. Thus, these rules serve as tests for whether or not a string of words is a constituent.

Exercise 3: For each of the following examples, test whether the underlined string of words is a constituent, using each of the constituency tests.

(7) The kittens slept at midnight.
(8) The kittens from next-door slept at midnight.
(9) The kittens from next-door slept at midnight.
(10) The kittens slept at midnight.
(11) The kittens slept at midnight.
(12) That guy from next-door baked a cake.
(13) Some people put those things on your bed.
(14) Some people put those things on your bed on the floor.
(15) Some people put those things on your bed.

Note on false negatives. In a perfect world, the constituency tests would have perfect validity. However, the world is not made to order, and it turns out that it's possible for constituents to fail one or more constituency tests. Thus, while (16) is generally true, (17) is not:

(16) If a string passes the constituency tests, then it is a constituent.
(17) If a string is a constituent, then it passes the constituency tests.

In other words, the failure of a string to pass a constituency test can be a *false negative* result. Constituents may fail at least some of constituency tests (i) because they are words rather than phrases, (ii) because they are finite, or (iii) because they are contained within so-called syntactic islands. [For example, see class discussion].

Argument 3: Ambiguity. Saying that sentences have internal structure helps us account for structural ambiguity. We can say that ambiguous sentences result when a string of words form constituents in more than one way.

Exercise 4: Account for the ambiguity of the following example by showing that the PS Rules assign two distinct trees for it. Specify which interpretation goes with which tree, and why.

(18) The detectives watched the man with binoculars.

Motivate your structures using the constituency tests (see class discussion).

Exercise 5: Provide tree structures for the following examples, and use the constituency tests to test those trees.

(19) Sara looked out the window.
(20) Sara looked up the number.
(21) Sara ran up the mountain.
(22) Sara rang up her friend.
(23) Sara threw the towel in.
(24) Sara threw the towel in the basket.

Exercise 6: Revise the PS Rules so that they generate the following sentences:

(25) The suspect will leave.
(26) The suspect has left.
(27) The suspect might leave.
(28) The suspect could cry.

(29) Those beds are soft.
(30) Those beds are very soft.
(31) The kittens like very soft beds.
(32) The kittens like soft beds.

===Note on coordination=====
Only identical constituents can be coordinated:

(33) a. I wrote to you and to him. PP Conj PP
b. I wrote a letter and a postcard. NP Conj NP
c. *I wrote to you and a postcard. PP Conj NP